Scope:

Step 1:

This project about global temperatures on country,city,state wise,to analyse this data we use aws s3,spark and AWS redshift data warehouse .

As par of this project ,first we load data to S3 and read data from there and apply spark on it to process the data and then load the data to aws data warehouse for querying required data.

Step 2:

There are 4 files include

* Global Average Land Temperature by Country (**GlobalLandTemperaturesByCountry.csv**)
* Global Average Land Temperature by State (**GlobalLandTemperaturesByState.csv**)
* Global Land Temperatures By Major City (**GlobalLandTemperaturesByMajorCity.csv**)
* Global Land Temperatures By City (**GlobalLandTemperaturesByCity.csv**)
* Attributes for **GlobalLandTemperaturesByCity.csv**)

Attributes for **GlobalLandTemperaturesByCountry.csv**

**Date – dt**

AverageTemperature

AverageTemperatureUncertainty

Country

Attribes of **GlobalLandTemperaturesByState.csv**

Dt

AverageTemperature

AverageTemperatureUncertainty

State

Country

**GlobalLandTemperaturesByMajorCity.csv**)

Dt